

学校编码: 10384

分类号_____密级_____

学号: 10220111152243

UDC _____

厦门大学

硕士学位论文

语义分布调查法在基础教育学习性词表研 制中的运用

The Use of Semantic Regimentation Distribution Survey
Method in the Word List Compilation for Basic Education
Learning

卢丹丹

指导教师姓名: 苏新春 教授

专 业 名 称: 语言学及应用语言学

论文提交日期: 2015 年 4 月

论文答辩时间: 2015 年 5 月

学位授予日期:

答辩委员会主席:

评 阅 人:

2015 年 4 月

厦门大学学位论文原创性声明

本人呈交的学位论文是本人在导师指导下,独立完成的研究成果。本人在论文写作中参考其他个人或集体已经发表的研究成果,均在文中以适当方式明确标明,并符合法律规范和《厦门大学研究生学术活动规范(试行)》。

另外,该学位论文为()课题(组)的研究成果,获得()课题(组)经费或实验室的资助,在()实验室完成。(请在以上括号内填写课题或课题组负责人或实验室名称,未有此项声明内容的,可以不作特别声明。)

声明人(签名):

年 月 日

厦门大学学位论文著作权使用声明

本人同意厦门大学根据《中华人民共和国学位条例暂行实施办法》等规定保留和使用此学位论文，并向主管部门或其指定机构送交学位论文（包括纸质版和电子版），允许学位论文进入厦门大学图书馆及其数据库被查阅、借阅。本人同意厦门大学将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于：

（ ） 1.经厦门大学保密委员会审查核定的保密学位论文，
于 年 月 日解密，解密后适用上述授权。

（ ） 2.不保密，适用上述授权。

（请在以上相应括号内打“√”或填上相应内容。保密学位论文应是已经厦门大学保密委员会审定过的学位论文，未经厦门大学保密委员会审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权。）

声明人（签名）：

年 月 日

内 容 摘 要

词表是教学大纲制定、教材编写、课堂教学和测试的主要参考依据之一。从 20 世纪以来,基于常用词和词汇教学的词表研制工作一直持续。但是一方面因为词表性质的差异,通用性词表和对外汉语教学用词表并不能直接作为母语教学用词表使用;另一方面,由于对基础教育教材词汇统计的研究方法和理论成果局限于对教材词语面貌和分布状况的反映,缺乏对词汇学习者主体认知发展的考虑,也不能够直接用以指导基础教育词汇学习大纲的编撰。因此,服务于基础教育阶段学习者的学习性词表迫在需求,而现阶段首先需要为其寻找出科学、合理的研制方法。

本文基于认知语言学、文化语言学和社会语言学的理论研究,从基础教育学习性词表的特点和性质入手,明确学生词汇习得的规律与学生认知发展、智力水平的密切关系,提出学习性词表需要收录基础性、语文性、阶段性和实用性强的词语,而目前以词频和分布数据为主的词表研制方法还不足以满足此要求。

本文进一步综合使用数据统计法、语义认知法和对比分析法,通过阐述基础教育阶段学生词汇学习与认知推进的密切关系,提出了为学习性词表的研制提出了结合词频、文本分布和语义义域色彩三个维度的语义分布调查法,运用《现代汉语分类词典》所建构的语义分类体系,以考察人教版小学语文教材词语语料库中词语语义分类情况为例,说明了语义分布调查法的运用过程、特点以及对学习性词表研制的作用和价值。

最后,本文从人教版教材词语语料库中选取按使用度位次由高到低排列的“前 3000 词词表”和在使用度统计法基础上运用语义分布调查法提取出的语义认知度由高到低排序的“前 3000 词词表”,从各个角度通过大量的数据和图表展示了两表的选词差异,对词语表的变化情况进行了较为细致的考察,旨在验证语义分布调查法的可行性、合理性,为基础教育学习性词表的研制工作提供参考和借鉴。

关键词: 基础教育学习性词表 研制方法 语义分布 认知发展

Abstract

The study of word statistics plays quite an important role in language teaching, dictionary compilation and information process of Chinese. While word lists are one of the main sources of setting the syllabuses, as well as teaching material construction, language teaching and testing. Based on the common words and words teaching, the study of word compilation continues since the 20th century. However, two major problems exist in word list compilation. On the one hand, because of the difference among the word lists' characteristics, the general word lists and the word lists for teaching Chinese as a foreign language can't be used directly in the basic education stage. On the other hand, the word statistics in the basic education textbooks have the limitations. Although they reflect the actual glossary constitutes of the textbooks, they are still lack of the investigations of the learners' cognitive ability, which can't guide the syllabuses set vocabulary requirements. Therefore, we need not only a word list for basic education learning, but also a scientific and pragmatic compilation method.

While the theories of cognitive psychology, cultural linguistics and sociolinguistics are adopted, the close relationship between the law of lexical acquisition and the students' cognitive ability is verified on the basis of the word lists' characteristics and nature. What's more, as for word selection, fundamental, hierarchical, functional and written language words are demanded to be taken into consideration, which the frequency and coverage statistics are not enough to meet compilation requirement of the word list for basic education learning

Combined the methods of database statistics, semantic cognitive and comparison, the progress of the vocabulary learning and cognitive development in the basic education stage is reconstructed as well. Furthermore, semantic regimentation distribution survey method under the word frequency, range, coverage

and semantic-regimentation is put forward for the word list compilation. The semantic-regimentation of the Chinese textbooks lexicon in the elementary school was depicted as a whole by the way of establishing the conjunction to “A Thesaurus of Modern Chinese” (TMC), basing on the language material database, in order to expound the process of how to use the method and its features and value as well.

This dissertation, from various angles, analyzes two word lists’ top 3000 words—— semantic-regimentation distribution word list and usage word list, which are picked up from the Chinese textbooks in elementary school language material database. So as to present the difference by lots of statistic data and charts, hence, semantic regimentation distribution survey method is proved more scientific and reasonable for the word list compilation and some reference can be provided for the word list for basic education learning as well.

Keywords: word list for basic education learning; compilation methods; semantic-regimentation distribution; cognitive ability

目 录

内 容 摘 要.....	I
ABSTRACT	II
第一章 绪论	1
一、课题来源.....	1
二、本文的研究对象、目标和方法	2
三、研究现状.....	4
第二章 基础教育学习性词表的性质	15
一、基础教育学习性词表的性质	15
二、基础教育学习性词表的基本特点	17
三、基础教育学习性词表的理论研制基础.....	18
第三章 基础教育学习性词表的研制方法	24
一、其他词汇统计方法	24
二、语义分布调查法的提出和运用	30
第四章 使用度排序前 6413 词的语义分布调查.....	42
一、词表与《分类词典》语义体系的匹配情况	42
二、语义义类体系层级分布调查	45
三、通过语义分布调查法提取词表	56
第五章 使用度词表和语义分布调查词表的对比分析.....	59
一、两个词表前 3000 词的总体变化	59
二、跌出词与新进词的相关数据对比分析.....	60
三、两个前 3000 词表相关数据统计	67
四、小结	72
第六章 结语	74

一、研究的结论.....	74
二、不足之处.....	75
三、下一步的研究工作.....	76
参考文献.....	77
附录 1 表格索引.....	82
附录 2 插图索引.....	83
致 谢.....	84

Content

Chapter 1 Introduction.	1
Section I Topic Origin	1
Section II Object, Objectives and Method	2
Section III Review	4
Chapter 2 Nature of Word List for Basic Education Learning . . .	15
Section I Nature of Word List for Basic Education Learning	15
Section II Basic characteristics of Word List for Basic Education Learning	17
Section III Basic Theories of Word List Compilation	18
Chapter 3 Compilation Methods of Word List for Basic Education Learning	24
Section I Other Word Statistics Methods	24
Section II Semantic Regimentation Distribution Survey Method	30
Chapter 4 Investigation on Semantic Regimentation Distributing of the Usage Word List's Top 6413 Words	42
Section I Study on Semantic Regimentation Distributing in the Word List by establishing the conjunction to TMC	42
Section II Hierarchical Investigation on Semantic Regimentation Distributing	45
Section III The Way to Use Semantic Regimentation Distribution Survey Method	56
Chapter 5 Comparison Between the Two Word Lists	59
Section I General Change Picture Between the Two Word Lists' Top 3000 Words	59
Section II Study on Difference between Two Word Lists' Unique Words . .	60
Section III Other Statistics Comparison Between the Two Word Lists. . . .	67
Section IV Conclusion for this Chapter.	72
Chapter 6 Conclusion	74
Bibliography	77
Appendix 1 Form Index	82

Appendix 2 Chart Index83
Acknowledgement84

厦门大学博硕士论文摘要库

第一章 绪论

一、课题来源

本文的研究来源于国家语委“十二五科研课题”——“基础教育学习性词表的研制”，由国家语言资源监测与研究中心教材中心承担。

（一）课题的提出

词汇是语言的三大要素之一，是语言技能的基础。无论母语学习抑或外语学习，词汇学习被普遍认为是语言学习的中心任务，掌握常用词更具有重要意义。“在词汇教学中，教师除了讲清楚词语的形、音、义外，还要讲清楚词语的用法。词语教学向下可延伸到语素，向上可延伸到词组和句子。学生掌握基本的语素构词规律和组词成句的规则都离不开词汇这一中心”^[1]。不仅如此，词汇知识与阅读理解、写作能力等基本语言技能更是密切相关。

反观国内，面对中小学生的母语教学，各地区教材之间词汇差异极大，教学成果的考核方式迥异，学习性词表的研制成为基础教育领域大纲制定者、教材编纂者、教学设计者以及制定学习策略的学生迫切的需求。

词表作为教学大纲的重要组成部分，既包含对现行基础教育状况的了解和学生学习认知特点的把握，还包含了对各科教材语言的整合和建构，对提高基础教育、学校素质教育的科学性具有重要意义。一个科学、合理的学习性词表为基础教育各领域翘首期盼，惟其如此，教材编写、词汇教学乃至课外读物的编写才能“有法可依”，基础教育才能真正走向更科学的道路

（二）研究意义

研制基础教育学习性词表的学术价值和实践价值在于：

1.能填补我国基础教育学习性词表的空白。在目前已经拥有了基本够用的通用性词表、对外汉语学习词表、专科性词表的情况下，再拥有基础教育学习性词表。恰恰是基础教育领域，对学习性词表有着更为迫切的需求。

2.能直接有助于基础教育各科教材编纂和词汇教学内容的落实，能使教学

大纲中有关词汇教学内容、要求得以具体化，解决基础教育各领域不协调的矛盾。

3.为服务于广大中小学生学习词典提供依据，也为编辑出版课外读物提供理论支撑，旨在更有效地促进课内外读物的互动，更有针对性地加强课外读物对词语教学的积极影响。

4.形成能反映我国中小学生词汇学习特点和规律的基础教育学习性词表的研制理论和科学方法。

（三）课题拟研究的主要问题、重点和难点

该课题拟解决的关键问题主要有三个，一是词表的量，即数量的多少；二是词表的分级，基础教育的学习性词表必须要根据学习阶段的不同要求解决词汇分级的问题；三是通用词表与学科词表的关系。通用词表反映的是中小学生学习词汇的一般能力，学科词表反映的是在各科知识学习中使用到的学科常用词语。

重点和难点主要有两个，一是学习性词表除了要具有一般社会词表的通用性、常用性外，还要能反映中小学生语言习得的认知需求和认知特点，这与社会一般用词有着明显的不同；二是学习性词表要能体现学习阶段的要求，能与学生的语言知识和语言运用能力，与学科知识的吸收和转化相匹配。

二、本文的研究对象、目标和方法

（一）研究对象

规定性词表有着明确的使用目的，它以反映语言系统中的词语为主，对其适用领域具有较强的指导性。学习性词表作为基础教育教学领域的规定性词表，其功能、要求在词表研制的过程中一直处于重要地位，并深深影响着词表的研制方法。本文的研究对象就是为服务于国内母语教学的词表确定其性质、研制的理论基础和研制方法。

1.通过词表研制的现状及数据比较来探寻基础教育学习性词表与社会通用性词表、对外汉语教学用词表在性质和功能上的差异。

2.突出学习者的认知需求、认知心理和认知特点，结合词汇教学的现状来形成学习性词表的研制理论。

3.本文主要研究的两个词表（语义分布调查词表和使用度表）依托国家语言资源监测与研究中心教育教材语言中心研制的“人教版小学 1-6 年级_分词总库”语料库，其语料来源于崔峦、蒯福棣主编的人民教育出版社的小学《语文》教材（初审时间为 2001-2004 年），共计 12 册。另外以商务印书馆出版的苏新春老师主编的《现代汉语分类词典》^[2]作为词语语义分类的参照标准。关于本文使用的词表的提取过程在第四章中将作详细介绍。

（二）研究目标

1.通过对研究语料中词语语义分布状况的调查，揭示学生词汇学习与认知发展的关系，并通过对教学大纲和课程标准中任务要求的分析，构拟出学生词汇学习与认知推进的过程。

2.提出计算词汇的语义认知判断标准，运用语义分布调查法提取出词表。

3.通过语义分布调查词表和使用度表之间的对比，揭示其存在的差异，并初步探讨产生差异的原因，验证运用语义分布调查法研制基础教育学习性词表更为科学合理的假设。

（三）研究方法

本研究在经过选取语料、数据对比、定量和定性结合的系统化集成处理过程后，提出基础教育学习性词表的性质、特点及研制方法。

1.频率统计法：通过 Access 数据库软件对语文教材词语的各个方面进行定量分析。语义类的分布比例确定语义类的重要程度，运用使用度公式和“语义认知系数”提取出两个词表，再使用数据库对这两个词表从多角度进行提取和分析，并对提取结果进行量化处理和统计分析。

2.语义认知法：从语文教材词语的语义分布特点来论证基础教育学习性词表独特的性质、理论研制基础以及语义分布调查法对学习性词表研制的作用和价值。

3.对比分析法：本文对使用度词表和语义分布调查词表进行了不同角度的

分析和对比，通过大量的数据和图表展示了两表的差异，对词语表的变化情况进行了较为细致的考察，逐步揭示语义认知度的加入使基础教育学习性词表的研制方法更加科学、合理。

三、研究现状

（一）国内词汇统计研究综述

词汇作为语言的建筑材料，是整个语言系统的基础，也是语言系统中最活跃、最具生命力的元素。词汇统计研究就是对词汇进行定量的研究，研究的成果就是词表。从手工统计到计算机操作，从卡片记录到以数据库为载体，词汇统计规模由百万级语料到千万级、再至亿级语料发展。词汇统计的相关研究在课程设置、语言教学、词典编纂、中文信息处理等诸多领域取得了丰硕成果。

1. 早期的词汇统计工作

现存最早对现代汉语词汇统计研究进行探讨的文章是 1922 年黎锦熙的《国语中基本语词的统计研究》一文，其中提出了对“国语”词汇进行统计研究的主张。^[3]

1925 年，陈鹤琴编著《语体文应用字汇》，对 12.7 万字的儿童用书、15.3 万字的报刊、9 万字的妇女杂志、5 万字的小学生课外作品、7 万字的古今小说、6 万字的其它类等——共计 554478 字量的白话语料进行了手工的字频统计，旨在确定小学生课本的用字量。这是我国最早针对母语教学领域进行的统计和分析，虽然只统计了字频，却开创了如何科学编制教科书研究的先声。^[4]

1930 年，民智书局出版了王文新的论著《小学分级字汇研究》，他以 91 万字的语料——包括当时的三种版本的国语教科书、各地的小学作文原稿以及儿童课外读物为研究对象，通过人工分词、手工统计的方法确定了小学识字的字量为 3799 字。这一研究是首次将词汇“统计”落实到实践中，让人们认识到常用字、常用词对于母语教育的重要价值。^[5]

1934 年，彭仁山对《三民主义用词统计与分析》^[6]一书进行了词汇统计，得出该书具有“不同的单词复词 4162 个”。

Degree papers are in the "[Xiamen University Electronic Theses and Dissertations Database](#)". Full texts are available in the following ways:

1. If your library is a CALIS member libraries, please log on <http://etd.calis.edu.cn/> and submit requests online, or consult the interlibrary loan department in your library.
2. For users of non-CALIS member libraries, please mail to etd@xmu.edu.cn for delivery details.

厦门大学博硕士论文摘要库